Evaluation of Artificial Intelligence Solution for FixYourMoney.com

May, 2022

Version: 1.1

Company: FixYourMoney.com
Industry Focus: Virtual debit card provider that adds credit card features
Established: 1st April 2020

Version History:

No.	Date	Comments
1.0	26-Apr-2022	Outlined the structure for the proposal
1.1	9-May-2022	Provided the AI areas and implementations for the company

Disclaimer: *The company, FixYourMoney.com, is a fictitious company used as an example to provide analysis towards study of AI Technologies as part of the degree course for M.Sc. in Artificial Technologies from University of Essex.*

1 Table of Contents

2 Introduction

FixYourMoney.com is a start-up fintech company that was started on 1st April 2020. The company provides credits and rewards to its customers as a banking platform based on customers' expenditures. It adds an analytics process to predict the benefits and rewards as a credit with actual credits provided.

3 Features

The banking application is rolled out as an app-only (IOS and Android) application as a day-to-day banking like a bank but bundling with customer centric experiences where the customers can:

- Subscribe and get an account with quick KYC and credit checks done
- Add all their banks into the app with the power of Open Banking and personalise expenses or let the app build an intelligent prediction
- Build their customised time-limited cards. Example, travelling to a country then no need to carry forex cards. Create a virtual card, add budget limits and expiry time.
- Get unlimited cashback. Example, bills paid from here gets you highest cashback

4 Information (Data)

Being a bank, following the FCA regulations (https://www.handbook.fca.org.uk/handbook/SYSC/9/1.html), the application requires to persist and audit records for its organisation, external providers and transactions associated by the applications involved. Below is the classification of important dataset that are being captured:

- Customer
    - Profile
    - Linked Banks
    - Wallet
    - Personalised Brands, Billing, etc.
    - Goals
    - Subscriptions & Cards
    - Rewards & Earnings
- Regulated Banks
    - Open Banking (OB) integrated
    - FCA approved
    - FixYourMoney.com subscribed
- Merchants (Brands, Billings, etc.)
    - FixYourMoney.com subscribed
    - Cashback value on expense
- Banking
    - Book (Money in, Money out, Money earned)
    - Transaction weightages
    - Transaction categorisation
    - Transactions flagged
- Virtual Cards
    - Card details (virtual number, provider, dates, CVV)
    - Card Provider
    - FixYourMoney.com subscribed
- Rewards
    - Merchants subscribed

  - Billing
  - Reward (applicable reward)
  - Reward awarded
- Customer services
  - Cases (Raised, Open, Rejected, Closed)
  - Notifications
  - General updates
  - GDPR
- Audits
  - Registration
  - Login
  - KYC
  - Profile updates
  - Banking Transactions
  - Virtual Cards
  - Credits
  - Customer services

5 Review

The review appreciates the key challenges the company faces that affects the customer conversions which cascades to the returns of the company.

The following key stakeholders were interviewed / workshopped:

- Product Owners & Analysts
- Marketing & Customer Services
- Compliance & Auditors
- Infrastructure & Security

to build a list of objectives supported with customer churn, conversions, and prioritized objective for the business (**business understanding**).

The objective is to build a list of business needs in the form of **business cases**. This requires a detailed analysis of the technology and problems faced by the company that directly or indirectly impacts customer conversions.

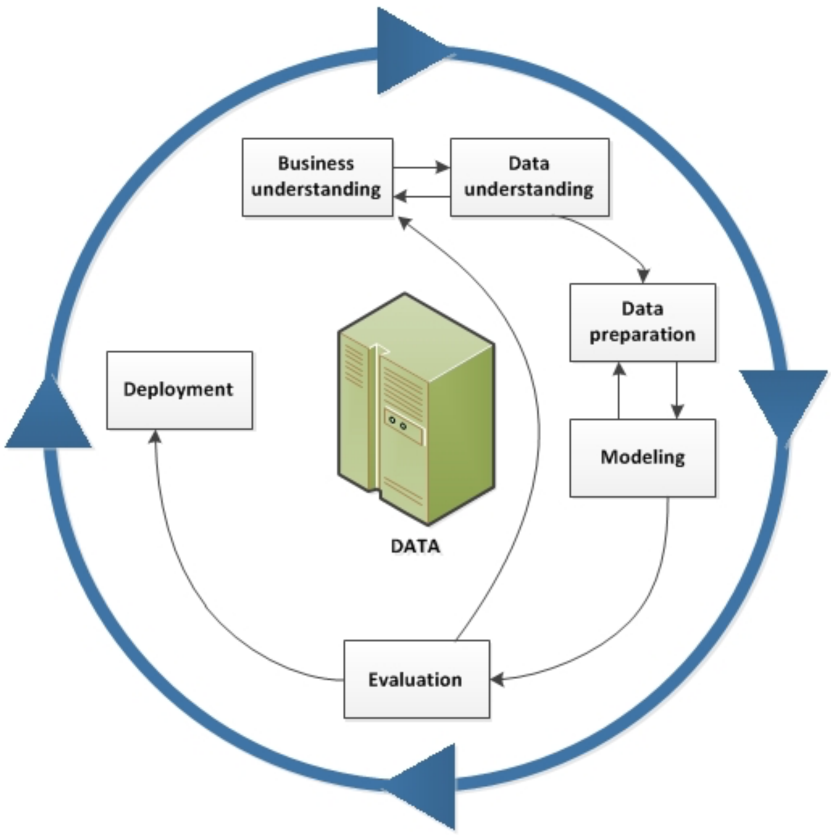

Figure 1: CRISP-DM (courtesy IBM)

The data and its flow need to be put under a lens to understand its spread and quality (**data understanding**). This step requires to interfere with the data more closely to make sure that chances of issues coming up in later stages are controlled and are low.

Data should be accessible i.e.

- Origin and sources of the data
- Does the **existing data** cover the business case?
- Are there any **external data** that are required to meet a case?
- Is there a need to create or capture **additional data** to achieve the requirement?

Once the data has been described (data quantity and data quality) and explored, it is then verified. Verification of data should lay out the data inconsistencies, for example, missing data, errors, incorrect metadata, etc.

Post the data understanding in conjunction with business, the key business cases were identified:

- Increase in brand subscriptions by customers
  - Results in near-to brand engagement
  - Results in customer buying linked with correct rewards
  - Lowers customer services engagement
  - Data is trained enabling cross-selling and promotions
- Chatbots as a customer service
  - Results in sharing load for customer services
  - Results in low customer drop
  - Results in resolving common Q&A
- Categorisation of transactions
  - Results in bucketing transactions for merchants
  - Results in training bots to serve comparisons

6 Prioritize

The expectation is that the outcome from the business and data understanding as a business case must align with the understanding of the key stakeholders. This is important because in further steps of the mining process there is time and money invested with resources (people, tools, etc.) that must measure with what the business agrees to.

The prioritized case is then put into the modeler tool to **prepare** and **model** the data. These steps take most of the project time. Here the data is – sampled, merged, sorted, aggregated, created with new attributes, and prepared into training and test datasets to build the model. **Modeling** is an iterative process where the model is fine-tuned by executing several models with different parameters and test data.

The output from modeling is a model must be an output that can relate to the business case. It should be measurable by comparing the results from more than one type of model. The important part of the assessment is that the model should be deployable.

The business case defines the **business criteria** and models should be **evaluated** with those criteria. Every model needs to go through the **deployment** step so that the business report provides the output based on the analysis. It does not matter whether the model evaluation was successful or not. If the business case is not met, then from evaluation the business understanding is triggered.

The business cases were then put into a 2x2 matrix between **business case** and **value**. The key case was: **Increase in brand subscriptions by customers.**

7 Resources

The resources are integral towards building an AI capability pool that can learn and build. It can be people, tools or processes. The assessment of existing strengths and weakness in relation to the business case brings out –

- Time to pilot the case
- AI and ML consultants required
- Working group to address data
- Tools – IBM SPSS Modeler for CRISP-DM, Azure ML Studio,

8 Solution

The object of the business is to enable the learning within the app to **"Increase in brand subscriptions by customers"**.

How does the learning help?
Customer subscribes → Brand subscription registered

Customer buys → Transaction registered → Apply filter (brand, category, price) → Apply filter (reward) → Customer is rewarded

The solution uses the vector regression outlined by *Giri, Chandadevi, et al, 2019* for fashion and apparel industry for classification of data to predict recommendations.

The application of filter during by the application can be trained to look into results from top brands used by customers, best rewards provided by brands, etc. for predicting similar purchases cashback returns by brands.

8.1 Approach (Using Azure ML)

Building of the model, as described in Information, Review and Prioritize sections are the building blocks to the approach for implementing the model in AI applications.

Based on the data analysis done (refer Information section), the data now is required to accessed and to be built as a chosen model. The activities required will uses the CRISP-DM methodology to get the chosen model (refer Review and Prioritize sections) along with the Azure ML Studio.

With Azure ML studio,

1. Raw data access
2. Pre-processing of the data
3. Prepares the data
4. Application of learning algorithm to the data
5. Create a Model and Evaluate it
6. Deploy the model on Azure
7. Provide model access to an Azure ML API

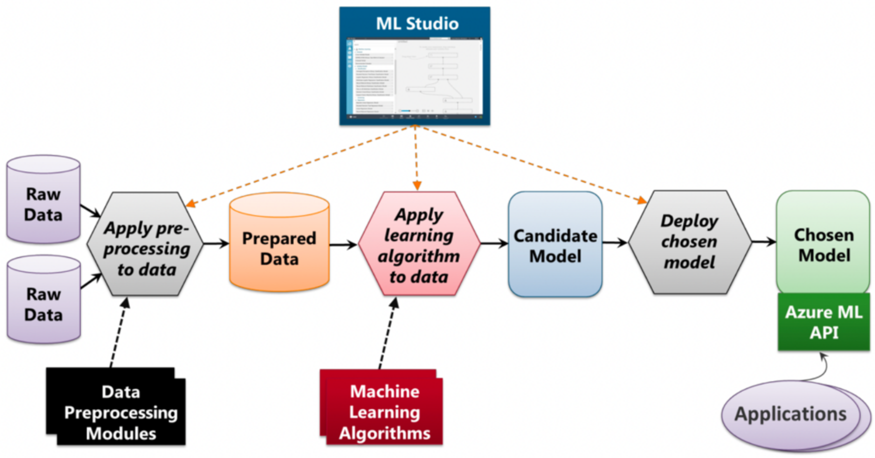

Figure 2: Azure ML data processing (Chappell & Associates)

The important part here is selecting the appropriate learning algorithm. The trials were done with **Boosted Decision Tree Regression** and **K-Means clustering** provided by Azure ML.

Decision tree regression is a supervised algorithm whereas clustering is an unsupervised learning.

The trials showed K-Means clustering as closer to the business case and hence this was chosen.

Once the flow is built within the studio and the model has been tested, it can be integrated with ML applications. Here the review and prioritisation of the data is important so that the application can predict as close as the business case.

This then can be integrated with likewise application to ingest the data and let it pass through the model for a result.

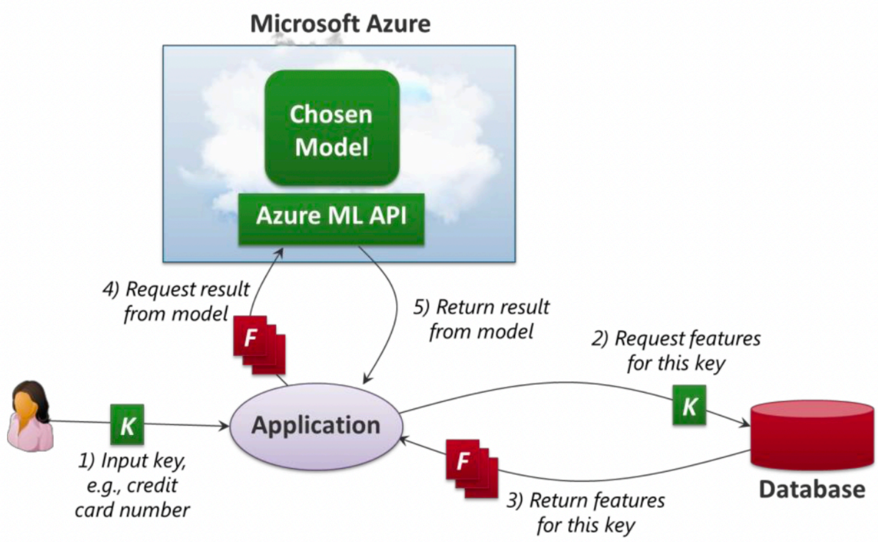

Figure 3: Integrated view for application processing (Chappell & Associates)

8.2 Deployments

Azure ML studio provides functionality to integrate the models directly with ML applications hosted within Azure or externals. Supporting synchronous and asynchronous integrations to provide predictions. This can be enabled for bulk or single requests.

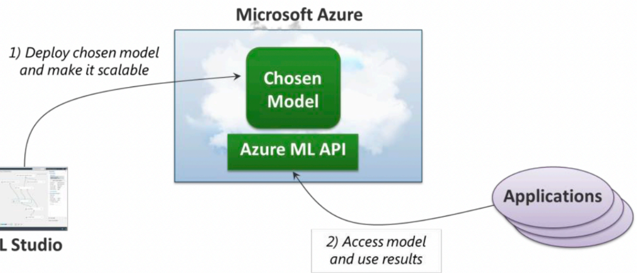

Figure 4: Model deployment to Azure Applications (Chappell & Associates)

9 Conclusion

With AI learning strengths of supervised and unsupervised learnings, it is still the requirement of data scientists that are pivotal in building the models from datasets that are consistent with resulting. This area needs to be evolved further to cater to any disruptions within the data ingestions.

Experiences from implementations should be documented to be reviewed for future cases and how this can be extended to other business cases like enhancing the customer services as virtual assistants.

The focus should be in identifying the best solution and not alternatives. It should have the capability to be enhanced for providing recommendations for required purposes by training the data.

The suggested algorithm may not be the only solution, but it is provided as an assessed solution based on the rigorous review process of the data sampled and tested.

The challenge and complexity arise in choosing the right algorithm. For this the business objective should be clear and model results should be reviewed with business outcomes because it could lead to regrettable waste in resources – time and money.

10 References

- FCA Handbook: https://www.handbook.fca.org.uk/handbook/SYSC/9/1.html
- IBM CRISP-DM: https://www.ibm.com/docs/en/spss-modeler/SaaS?topic=dm-crisp-help-overview
- CRISP-DM: https://www.datascience-pm.com/crisp-dm-2/
- Azure ML: https://azure.microsoft.com/en-gb/services/machine-learning/#features
- Boosted Decision Tree Regression: https://docs.microsoft.com/en-us/azure/machine-learning/component-reference/boosted-decision-tree-regression
- Giri, Chandadevi, et al. "A detailed review of artificial intelligence applied in the fashion and apparel industry." IEEE Access 7 (2019): 95376-95396.
- Bhimani, Janki, Miriam Leeser, and Ningfang Mi. "Accelerating K-Means clustering with parallel implementations and GPU computing." 2015 IEEE High Performance Extreme Computing Conference (HPEC). IEEE, 2015.